université angers

13, allée François Mitterrand
BP 13633
49100 ANGERS Cedex 01
Tél. : +33 (0) 2 41 96 21 06

Web : http://www.univ-angers.fr/granem

Granem
Groupe de Recherche ANgevin en Economie et Management
UMR-MA n°49

Who cares? Europeans' attitudes towards the disclosure of personal identity data

Caroline Lancelot Miltgen
GRANEM, Université d'Angers

Margherita Bacigalupo
European Commission Joint Research Centre

Wainer Lusoli
European Commission Joint Research Centre

Janvier 2011

Document de travail du GRANEM n° 2011-01-026

AGRO CAMPUS OUEST

Who cares? Europeans' attitudes towards the disclosure of personal identity data

Caroline Lancelot Miltgen, Margherita Bacigalupo, Wainer Lusoli

Document de travail du GRANEM n° 2010-04-026
Janvier 2011
Classification JEL : M14, M15, M38.

Mots-clés : Dévoilement de données personnelles, Respect de la vie privée, réunions de groupe, analyse textuelle, différences culturelles, écart de génération
Keywords: Personal Identity data disclosure, privacy, focus groups, textual analysis, cultural differences, generation divide

Résumé : Dans les milieux politiques, une attention croissante est portée sur les questions de sécurité et de vie privée liée à la divulgation de données personnelles. Le dévoilement de son identité semble être un élément intégral du style de vie des jeunes. Par ailleurs, il semble y avoir des différences d'attitudes entre les jeunes adultes et les cohortes plus matures. Cette recherche s'intéresse aux pratiques des citoyens en ce qui concerne plusieurs catégories de services en ligne. Il cherche a analyser les opinions, attitudes et motivations envers le management de données personnelles et étudie les préférences déclarées en matière de régulation. L'étude a été menée dans 7 pays européens répartis dans 4 régions différentes où les attitudes envers la vie privée diffèrent. Deux réunions de groupe ont été menées dans chaque pays, un ciblé sur les jeunes de 15 à 25 ans et l'autre sur les adultes. Deux types d'analyses textuelles complémentaires ont été menés. L'analyse de contenu permet d'identifier les principaux thèmes abordés par les participants tandis que l'analyse discriminante permet d'identifier les similarités et différences au sein des discours par rapport à des variables telles que l'âge des participants et le pays d'origine. Même si cette recherche n'a pas pour but de donner une vision complète des perceptions des Européens sur ce sujet, elle permet de souligner des différences significatives dans ces perceptions, que ce soit d'un point de vue générationnel (écart entre les jeunes et les adultes ou culturel. L'étude a donc des implications scientifiques et politiques importantes. Les résultats doivent aider à bâtir un questionnaire visant à être administré dans les 27 pays européens sur le même thème. Les conclusions aideront également la Commission Européenne à définir son agenda à venir sur les questions de Société de l'Information.

Abstract: In policy circles, there is increasing attention to the privacy and safety of young people's personal identity data. Identity disclosure seems to be an integral part of young people's lifestyles. Also, there is mixed evidence on the different attitudes regarding disclosure between young adults and more mature cohorts. The present research examines peoples' practices in relation to various services (e.g. social networking, health). It gauges opinions, thoughts and motivations towards personal identity data management and covers policy preferences in relation to the protection of identity data. The study was conducted in seven EU Member States, covering four regional groups where attitudes to disclosure have been shown to differ. Two focus groups were run in each country; one with young people aged 15 to 25, the other involving adults (25 to 70 years old). Analysis is based on two complementary qualitative techniques, using textual analysis software. Content analysis was used to identify the main topics emerging from the groups' interactions, while a 'discriminate' analysis was performed to obtain a deeper insight into discourses' similarities and differences in relation to specific variables such as age and regions/countries. While this research does not aim to provide a comprehensive view of Europeans' perceptions, it highlights significant differences, particularly between young people and adults and between the different countries (cultural differences). This study consequently has significant scientific and policy implications. The results will help shape a final questionnaire for a EU27 survey on the same subject. This work will help the European Commission direct the Information Society agenda in the years to come.

Caroline Lancelot Miltgen
Faculté de Droit, Economie et Gestion
Université d'Angers
caroline.miltgen@univ-angers.fr
Margherita Bacigalupo
European Commission Joint Research Centre
Wainer Lusoli
European Commission Joint Research Centre

This study was funded by the European Commission IPTS (Institute for Prospective Technological Studies) Joint Research Centre: EC JRC Contract IPTS n° 151592-2009 A08-FR (December 2009 to April 2010, budget: 49000 Euros) with Caroline Lancelot Miltgen as the research project leader at the University of Angers (France).
The authors want to thank Dominique Peyrat for her help with the Wordmapper© analysis.

1. Introduction

Personal identity data management (PIDM) and especially disclosure is central to a significant number of policy areas including growth and competitiveness, data protection, privacy and safety. At the same time, as Europe is in the process of changing from a network society to a knowledge society, policy makers need to better identify and measure the impact of technologies in order to help shape society as a whole and the 'digital' lives of European citizens in particular. The disclosure of personal data is also central for companies which want to know a lot about consumers' identity, tastes and preferences in order to propose personalized offers which will have a better chance to increase sales.

People's motivations and fears to adopt technologies, digital innovations and new online services are well known. EU citizens, for example, are greatly concerned about the security and protection (i.e. confidentiality) of the personal identity data that they disclose in exchange for e-services. However, across member states, there is significant variance and despite an existing strong protection offered by the Legislation (e.g. Directive 95/46/CE, for the protection of personal data). Citizens also feel increasingly responsible for their own security and privacy but are slow to adopt security and privacy enhancing technologies. This is one consequence of the famous privacy paradox described in numerous publications about privacy. All these gaps and contradictions hamper a clear understanding of people's views and impede the full adoption of new identification technologies (such as e.g. biometrics or RFID) and the development of a fully protective policy framework.

The literature on privacy and personal identity data self-disclosure mostly focus on the impact of individual factors on privacy concern (e.g., Campbell 1997; Milberg et al. 1995; Phelps et al. 2000) and then on the influence of privacy concern on subsequent behaviours, such as transacting, self-disclosing, or giving false data (e.g., Dinev & Hart 2005, 2006; Lwin et al. 2007; Changi Nam et al. 2006; Hong Sheng et al. 2008). Despite this literature however, we still don't really know how people make the decision to disclose (or not) personal data to public or private entities. We don't know either why they sometimes give data that seem sensitive (e.g. data about health or financial data) to peers (in Social Networking Sites) or to companies and why protection tools against possible misuse of these kinds of data are not applied.

Further research with the aim of fully understanding the personal data protection issue in Europe is now required in a view of 1/ implementing policy measures that could address it directly and 2/ designing successful and protecting future electronic identification systems. The purpose of this paper is thus to present the results of an exploratory study of Europeans practices, attitudes and policy preferences with regards to personal data identity disclosure, with a view of directly assisting policy formulation and consensus in this area.

The research consisted of the design, organisation, set up, analysis and reporting of focus groups on EU citizen's attitudes and behaviours regarding personal identity data management. The aims of this qualitative study are threefold:

- Perform a dry-run of a previous questionnaire in order to assess its suitability, fit and focus before carrying out a Pan-European survey on this subject
- Applications, problems and issues assessments around the PIDM topic in order to complete and revise the questionnaire
- Provide evidence as to EU citizen's understanding, motivations, fears, perceived risks and corresponding behaviours in relation to PIDM and policy protection

As a whole, the aim was to provide substantial advice on how to conduct a EU27 Special Euro barometer survey on this topic.

2. Foundations and purposes of the research

Culture plays an important role in how people interpret information. It is true also of how users perceive security, privacy or PIDM-related issues, such as feeling secure, feeling private, and feeling trustful. Culture has been shown to differ, even in countries in the same geographical area. This is true also in Europe (EU27) where culture is totally different in

countries from the North, the South, the West and the East. For example, trusting someone or something is a wholly different matter in Finland, than it is in France. Culture can even be different in countries that are very close in a geographically (i.e. Spain and Italy).

As the actual European regulatory framework is the same for all the 27 countries and as privacy is a very personal and cultural matter, it is thus very important to better understand the different views (motivations, fears and perceived risks in self-disclosing personal identity data) of people coming from different geographical and cultural areas in Europe even if all these people pertain to a same geographical 'Block' known as EU27. The way in which people from different EU27 countries 1/ disclose personal identity data to public and/or private entities and 2/ judge the efficacy of the actual policy framework may explain differences in privacy concerns and thus in the adoption of future electronic identification systems and protection means. It is thus important to measure and explain these cultural differences and this is one of the aims of this paper: to present the perceptions about identity data disclosure and privacy from people coming from different areas in Europe.

The generation divide that exists as regards to the relation people have with technologies is also well recognised. In the literature, we consistently find that the under 25's stand out as different from older age groups in their interests, values, attitudes and even behaviours relative to new technologies. For example, studies on identity show that 15-25 year olds are intensely interested in their personal identity : they don't know what they will turn out like and they 'try on' various personas. At this age, young people are discovering and developing their skills and talents. Many of them are high consumers of e-technologies: 88% of the 16-24 year olds of EU27 are connected to the Internet versus 60% of all the EU27 individuals (Data from Eurostat, 2009). The young people are consequently individuals whose practices on line are often in advance compared to the average European Internet surfers and to adult people. Moreover, they represent possible opinion leaders in the area of the Internet. They are also inclined to communicate their opinions by using new technological tools such as virtual communities, buzz communication and blogs.

Most literature on the PIDM and privacy topics focuses on a specific age category, whether pre-adolescent (e.g. Lwin et al. 2008 adolescents (e.g. Montgomery 2000 or adults (e.g. Hoadley et al.). Young people however, have specific uses of ICT (e.g. social networking) and specific interests/motivations (e.g. videogames, entertainment) that are not so important for older people. Thus, young people's motivations are very different from the motivations of adults. The comparison of young peoples' perceptions with that of adults with regards to PIDM and privacy issues is indeed important, in order to get some prospective insights about the needs/attitudes and behaviours of all the population and especially of people who will use these technologies in the future. Knowing the specific concerns of youngsters and adults will lead to information about whether public and private organisations have to adopt public policies and technologies targeted to the specific needs of each category.

3. Research method

3.1 The choice of discussion groups

Qualitative research encompasses a variety of methods that can be applied in a flexible manner to enable participants to reflect upon and express their views about the subject. In our case, a group discussion method whose main characteristics are presented in Table 3.1 was used. This consisted of groups of about ten (i.e. eight to twelve) participants attending a ninety minute discussion - conducted by an experienced interviewer - concerning the provision of relevant stimulus material (i.e. the interview guide with main topics to discuss).

The main purpose of this type of discussion is to gain insights by creating a forum where participants feel sufficiently relaxed to reflect and portray their feelings and behaviour using their own vocabulary, language and logic. This technique was particularly valuable here as a free-flowing conversation that is respectful and not condescending to participants appeared, although people with social and cultural differences were mixed.

Table 3.1 The main characteristics of the group discussions

Criteria	Characteristics
Purpose	Exploration and description
Sample size of the groups	Eight to twelve
Physical setting	Relaxed and informal atmosphere
Duration	Ninety minutes
Interviewing	Logical sequence (following the guide)
Recording	Use of video recording
Transcripts	Necessary
Key benefit	Group members feed off each other and creatively reveal ideas that the researcher may not have thought of
Key drawback	Group members may feel intimidated, shy or may not reveal anything

3.2 The interview guide and the issues explored

Discussion was structured using an interview guide that probed answers to 3 topics[1]:

Topic 1: practices and attitudes with regards to personal identity data disclosure
Topic 2: responsibilities, policy knowledge and preferences
Topic 3: perceptions about proposed scenarios

Following the interview guide, the interviewer initially revealed to the participants that the discussion was going to be about "new technologies and systems that permit the monitoring and processing of your personal identity data". The group discussions were recorded using digital audio and video recording equipment with the consent of all participants.

A key strategy employed in these discussions was to explore how, and in what ways, people deliberate on issues of identity, privacy, security, control, protection, responsibilities and policy preferences. By studying the process through which people make the decision to disclose or not personal data to specific public or private partner, we were able to shed light on how public perceptions are formed and maintained, in particular the reasoning that people use to make sense of regulation systems and control in this area of research.

This method enabled us to investigate:

1. How people consider their identity and the delivery of personal data in general.
2. The extent to which people accept to disclose personal identity data, in which conditions (circumstances, motives), to whom and how far this may outweigh perceived costs (risks) in terms of privacy and confidentiality (issues of trust, privacy and value).
3. The issues of control, protection, responsibilities and redress
4. The knowledge people have about the regulatory framework, their policy preferences and the perceived gaps with reality

3.3 The 'sampling' of participants

As culture can even be different in countries that are geographically close, it is important to choose the countries in which to conduct the focus groups with scientific rigor. As for timing and budgetary reasons studying all the countries in the EU27 is not feasible, it seems best to focus on a target population that reflects practices and views in several European regional blocks.

[1] Appendix 1 presents some examples of the issues and questions covered with the participants in each topic discussed. The whole interview guide is available upon request to the author.

Four regional blocks based primarily on their geographical area were identified for the purpose:
- Block 1: Denmark, Finland, Sweden, Estonia, Latvia, Lithuania
- Block 2: Bulgaria, the Czech Republic, Hungary, Poland, Romania, Slovakia, Slovenia
- Block 3: Austria, Belgium, France, Germany, Ireland, Luxembourg, the Netherlands, UK
- Block 4: Cyprus, Italy, Greece, Portugal, Spain, Malta

The countries for the focus groups were then chosen in relation to two relevant conditions:
- Surveying two counties in each block (except unknown issues)
- Choosing these countries so they are representative of the possible differences within the block (i.e the other countries in that block) and between all blocks

This representativeness is considered in relation to 3 elements (see Appendix 2):

- First, the general ICT development level of these countries which will be considered using the 3 sub-criteria presented hereafter:
  - Internet access: proportion of households in the country who access the Internet and/or who have an Internet access at home;
  - Uses of the Internet for online shopping;
  - E-services development in relation to public services (level of e-government initiatives and use of such e-gov services).
- Second, the socio-economic development level of the countries (see Appendix 2)
- Third, their "place" in Europe (see Appendix 2 for deeper understanding)

Considering these elements, for the first regional block, we propose to survey one "Nordic members" states such as Finland, Sweden and/or Denmark[2] as there is a vast majority of households who have access to the Internet as well as an important use of e-gov services in these countries. Estonia is also an interesting country as the e-gov online availability is quite high but with figures about access and uses not as high. Moreover, Estonia is very different from North EU27 countries with regard to the countries' place in Europe and socio-economic development.

Poland and/or Romania (from Block 2, East Europe) can be interestingly compared to Estonia (Block 1) as they are from the east of Europe but with quite different level of ICT development. Poland for example has a higher level of proportion of households buying online. Romania on the contrary is one of the only EU countries where a higher proportion of households use a dial-up connection as well as the lowest proportion of people with e-skills.

Germany (Block 3) is also an interesting country in that the figures are just below those of the "Nordic members" but having a different cultural and geographic background. France will be interesting to compare to Germany as it is moderately developed with regards to ICT but has been one of the most dynamic countries in the EU27 in terms of netsurfers between 2007 and 2008. Spain is quite similar to France in terms of ICT development but with different cultural and geographical backgrounds and a different progression rate. Greece, also a country from the South, will be interesting to compare to Spain as the figures for ICT development are largely below those of Spain.

As the internet is the basis for this study it was important to use a sample of countries that included a large distribution in this category. So, two countries chosen for their low internet use: Romania (30%) and Greece (31%). Two countries were also chosen for their high internet use: (Denmark[3] (82%) and Germany (75%)), with France (62%), Estonia (58%), Spain (51%) and Poland (48%) having percentages in between the high and low internet use groups. These countries represent the European dispersion of internet use well, their average being 54.6% which is close to the European average of 57.2% (2.6% difference). Correlations were also found for Internet use with e-commerce and e-government availability.

[2] Finally, no country from this geographical area participated in the study due to time constraints.
[3] Denmark that was initially retained finally dropped out.

Finally, the proposed countries have a relatively equal geographical repartition throughout Europe: two countries in each of the four cardinal directions.

As generation can also potentially divide the relation to new technologies, we consequently proposed that one focus group was conducted on young people aged 15-25 (before 15 the children won't be really at ease with these discussions and after 25 they will be considered as adults) and the other on the adults (from 25 to 70 years old). We thus conducted 2 focus groups in each country for a total of 14 focus groups in 7 countries.

This choice should permit the focus to be on the discussion and thus allow useful analyses of the topics of specific interest to the target (e.g. leisure for the young generation and health for the older one). However, there is also a large diversity within each generation with regards to for example gender, "professional status" and geographical location. Moreover, these factors could affect people's attitudes and behaviours towards PIDM (Aslanidou & Menexes 2008). For example, studies on perceptions about ICT, Internet or digital identity all show that gender has a significant influence on those perceptions (M. Tsai & C. Tsai 2010, Broos & Roe, Hoadley et al., Papastergiou & Solomonidou 2005). We consequently propose that the choice of participants within each focus group is also – as far as possible – constrained to respect some diversity in terms of age, gender and 'professional status'. For example, as there were around 10 participants in each focus group, we proposed to have half men and women, 1/3 before 18 and after 21 and 1/3 of students in the young focus group[4]. These indications have been given to each instructor in order for him/her to find people in each category. Indeed, if we didn't standardise the composition of the groups, it would have been difficult to say if one difference between two groups was due to the difference in the composition of the groups or to other elements.

3.4 The course of the study and the characteristics of participants

All the discussions were run in the native language of both the participants and the interviewers. The advantage of having country-of-origin speaking interviewers was considered important, in order to facilitate discussing such delicate research topics as "personal identity data" or "privacy". In order to create a trusting atmosphere during the interviews, having native-born people as conversation partners is likely to have enhanced the level of detail of the discussions with the interviewees.

As the discussions were run in different languages and with different interviewers, we briefed them quite intensively to ensure that the discussions were run in the same manner. We also gave them precise instructions to follow, in order to moderate the discussions. For example we remembered to respect the participants by providing anonymity, not deceiving them and conducting discussion in a way not to cause embarrassment or harm.

To inform the participants about these conditions and ensure that they were respected, they received a "consent for voluntary participation" form that they had to sign. They were also given 20€ as an incentive for taking part in the study. At the end of the discussion, the interviewees were asked to fill in a short form which comprised questions on areas that seemed relevant for investigating PIDM such as demographics, use of the Internet and email, Internet skills and relation to privacy. This information offers the opportunity to examine the differences in participants' answers in relation to their profile. Two of these variables were used for the purpose of this study: country-of-origin (or nationality) and age.

Around twenty participants (half youngsters and half adults) finally took part in the discussions in each country (e.g. 21 in Germany and 19 in Spain) for a total of 139 participants[5], of which 65 (46.8 percent) were male. 45 people interviewed are students and 28 are salaried. A high number of participants (61) are between 20 and 30 years old and 27 are less than 20. Most of the participants check their emails several times a day (45.9 per cent) or once a day (28.1 percent), with a many perceiving themselves as having a high (50.4 percent) or medium (23.3 percent) level of Internet skills on a 1 to 7 Likert scale.

[4] These instructions are wrapped up in Appendix 3 along with a summary of the whole sampling characteristics.
[5] The main characteristics of the participants are provided in Appendix 4.

3.5 Data analysis methods

We chose to run two different kinds of qualitative data analyses using dedicated software packages - Alceste© for the content analysis and Wordmapper© for the discriminate analysis - which permitted a more objective description of the issues discussed by the participants (in the first case) and a more objective measure of the differences in views and concerns according to specific variables such as nationality and age in the second case.

Qualitative data analysis packages do not automate the analysis process; nor is it their purpose. The process of coding for example always depends upon the interpretations made by the researcher. The overall results and theories that emerge from the analysis also depend upon interpretations made by the researcher. Both can be conducted manually, but by using software, the researcher can manipulate the data far more efficiently to help identify patterns, interconnections and ultimately to develop or test theory.

3.5.1 Content analysis

We first ran a content analysis in order to determine the topics most referred to by the participants on the theme of e-identity. Content analysis is a classical research tool to describe qualitative data which enables the researcher to make sense of large amounts of textual information very quickly, identifying its main properties, e.g. the frequencies of most used keywords. The methodology is grounded in Statistical Textual Analysis itself based on lexicometrics. The basic hypothesis is that language levels and text structure can be inferred from recurrent distributions of words.

The use of statistical textual analysis offers an extremely rich exploratory approach, both for the comparative study of texts and for the understanding of their content. This type of application is well-established and its effectiveness is widely acknowledged (Lebart and Salem, 1988, 1994). The textual statistics are a valuable tool for the reading and comparison of the transcriptions. Their application to the field of marketing and IT research emphasises their potential and the results that can be obtained (Gauzente and Peyrat-Guillard, 2007).

For this content analysis, we used a French software package called Alceste©. Its aim is to quantify a text to extract its strongest significant structures so as to draw the essential information contained in the textual data. Research has shown (Benzecri, 1980 and 1981; Reinert, 1986) that these structures are closely linked to the distribution of the words in a text and that this distribution is rarely done at random. The main function of Alceste© is to describe, classify, assimilate and automatically synthesise the text. The program works by applying a set of statistical clustering techniques to the text. Different forms of language relating to the research topic are identified by the program, and the text is categorised into clusters according to the distribution of the vocabulary found. The main method used by Alceste© is the Descending Hierarchical Classification (HDC). This method carries out successive splits of the text. It finds the strongest vocabulary oppositions in the text and then extracts some categories of representative terms. In reality, Alceste© performs two classifications to avoid any influence due to the text being split, and to guarantee stability.

3.5.2 Discriminate analysis

To reinforce and complement this content analysis, a second type of analysis (i.e. a “discriminate” analysis) has been run using other specific qualitative data analysis software. The WordMapper© software offers a complementary view of the corpus by running two kinds of analyses: Correspondence Factor Analysis (CFA) and cross tabulation. Correspondence Factor Analysis is a multivariate technique developed by Benzecri (1981) which detects possible associations and oppositions existing between variables (such as nationality or age in our case) and words. The projection of the variables (corresponding to participants’ profiles) and the words onto the same set of factorial axes enables two-dimensional graphs to be drawn which offer aid in the interpretation of the results. Cross tabulations are complementary as they outline the positive and negative specificities of each part of the

corpus. They give the words most frequently used by the participants depending on the occurrences of the variables studied i.e. age category and country of origin in our case.

4. Presentation of results

4.1 Ensuring the quality of the results

The internal validity of the results were ensured by following Yin's instructions (1989), who advised the comparison between the empirical results highlighted in the qualitative study and the theoretical proposals resulting from the literature. Our results were thus compared with the literature, to corroborate some of the results found in this qualitative phase and to disconfirm other findings.

The external validity generally constitutes a drawback for qualitative research. In our study, except to show that the choice of the participants would have been badly reasoned, the diversity of the people interviewed (almost 140 interviewees of both gender and all ages) is ensured. These people indeed come from countries that were scientifically chosen as being representative of the EU27 diversity on the basis of three main criteria (ICT and socio-economic development levels and place in Europe). As this choice meets the Mason's (1996) concept of being 'theoretically generalisable' (with all the necessary precautions), the findings presented here should thus be taken to indicate current European (i.e. French, German, Greek, Spanish, Polish, Romanian, Estonian) attitudes towards personal identity data management, although not a quantitative distribution of those attitudes.

To ensure the reliability of the research, we used two different methods. The first consisted of running the discussions by complying with the elementary rules of interviewing (kindness, empathy, involvement, and sensitivity). This makes it possible to ensure that the data obtained really correspond to the thoughts of the people interviewed. Moreover, we ensured that all the interviewers used the same techniques and maintained a high quality in their interviewing. All of them received detailed instructions on how to conduct the discussions and select the participants, as well as general instructions about interviewing. The triangulation is also a technique usually employed by qualitative researchers to consolidate the credibility of the results (e.g. Ammenwerth et al. 2003; Barnes & Vidgen 2006 Kaplan & Duchon 1988). We used both theoretical and methodological triangulations, the first one by testing various theories allowing for interpretation of the results and the second by using two kinds of analysis in order to compare the results and ensure the validity of the conclusions.

4.2 Content analysis: people's perceptions about PIDM and regulatory issues

The four main topics that were found through the content analysis are presented bellow, providing for each thematic area a significant excerpt of the FG participants' discourses.

The first topic related to Personal Identity Data Disclosure (PIDD) and use. Although most participants find the personal data collection quite intrusive - and then declare lying to obtain *pseudo-anonymity* – they sometimes consider the data disclosure as a quasi *compulsory* act in order to obtain the e-service required in exchange. Some participants also recognize that data disclosure may permit some *benefits* (e.g. promotion leaflets) and that people shouldn't always see it as a *constraint* and imagine undesirable consequences. Other respondents consider PIDM as a *trade-off* between both constraints/risks and benefits. However, all respondents advise disclosing only *insensitive* data to people/organisations you know/trust.

Table 4.1 Topic 1: Personal Identity Data (PID) Disclosure and Use

Participants reactions re PID Disclosure and Use	
Pseudo - Anonymity	'But the only thing necessary for that is our email. What do they need our name and surname for? So, why not give fictitious ones and a real email?'
Compulsion	'One hasn't really got a choice. For example when pursuing some goal such as getting a new email address or a new account. One is basically forced to do it. Well not all information, but the most important data. And you have to disclose it. Yes, they do charge a minimal fee.'
Benefit	'You disclose your data because no more than a name and an address is required, that is used by companies to send you promotion leaflets and information staff, which are often for your benefit'.
Constraints	'I easily give my name, first name and my address. When we must use internet I do not ask myself all these questions. It s paranoia to think that it will be stolen.'
Trade-off	'It must be that one decides according to the circumstances, since if one doesn't disclose any data, then I can just as well stop using the internet. It already starts in order for me to obtain an email address, I have to divulge it. It will be somewhere in between. Not revealing anything won' t work.'
Sensitivity	'It depends on the data you disclose; it depends on how close to me they are and on how private and secure this data is.'

Participants also highly discuss the issues of privacy and control. Most respondents consider disclosing personal data online as *loosing control* and some think it can go even towards a *breach of privacy*. A large proportion of the participants are concerned by such intrusion as they consider the risks of data *misuse* as very high, these risks being both difficult to anticipate and largely a future threat / not an immediate threat.

Table 4.2 Topic 2: Privacy and Control

Participants reactions re Privacy & Control	
Function creep	'How can I be sure that this data will only be used for this purpose and no other?'
Loss of control	'As soon as you put information on the internet you lose control of what you have and you no longer control anything in fact.'
Privacy breach	'But I believe that we re going more and more towards the breach of privacy of people, and we re going more and more towards dictatorship.'
Misuse	'I mean, these risks are difficult to anticipate. Because, the risk lies in improper use of this data. A bank won't use it against us, but this data may leak out from the bank s database. Someone may take that data away, and we can t anticipate what use this person will make of it.'
Future risks	'But I think that Facebook will hurt the actual generation for work later. They really displayed and this will be prejudicial. Secondary school pupils don't realize that there can be a potential danger.'

As regards to protection and regulation, people clearly express a need for more efficient and secure *regulation* and feel sorry for such *power imbalance*. However, we notice an evolving distrust in public authorities which is even reinforced by the feeling that there is *no miracle protection/security* on the Internet. People don't really know how they can make sure that their rights are respected (*redress*) and thus often use *self-protective* measures such as not registering online.

Table 4.3 Topic 3: Protection and Regulation

Participants reactions re Protection and Regulation	
Need for regulation	'My expectation would actually be that there are very clear legal guidelines for the use of such surveillance cameras.'
Power imbalance	'That's the problem, they're always legally covered. Lawyers aren't stupid. That's why they write these endless pages because they want to insure themselves against everything. But if protection were greater, we probably wouldn't be sitting there.'
Evolving (mis)trust	'But the trust in the authorities is diminishing as well, since the tax card and tax return will be eventually processed electronically. The trend is more and more towards being done online and with that the trust disappears as well. i don' t know, does the data go there or does it take a detour or the like.'
No security	'Anyway there's no miracle protection on the internet. You can use all the firewalls and anti viruses you like, it doesn't do any good.'
Redress	'I' m thinking of filing a complaint next time. If they ask me for my phone number again, I'll say I am not giving it because your phone calls are causing a disturbance. What right do I have?'
Self-protection	'Not to register is the best protection.'

In people's views, the responsibility of data misuse is mainly shared within three actors: the respondent *him/herself*, the *companies* collecting and handling the data and the *state* which should protect its citizens.

Table 4.4 Topic 4: Responsibility

Participants reactions re Responsibility	
Self	'It's up to each person to say I'm going to make my Facebook private, I' m going to post this photo, I want to send this photo to everyone and it's up to each person to be responsible.'
Companies	'Because the provider is responsible and not the state. the state can' t always be everywhere and say don't do that'
State	'I think everyone is responsible themselves. But in a way the state should protect your data, once they're made public. But if you disclose them in the internet yourself, then it's also your own fault. Then the state can't do anything about it.'

These general patterns can however be different considering the people's country of origin or age. For the purpose of this paper, we ran a country analysis followed by an age analysis to study whether there was or not any cultural and/or generation divide with regards to people's attitudes and behaviours towards PIDM and privacy concerns.

4.3 Country analysis results: any cultural divide?

The table below presents the themes/classes that Alceste© found per each country which correspond to the main topics/issues discussed by the focus group (FG) participants. The discourses focused on only three main topics in Poland and were more diverse in

Estonia, with six topics. A colour code is used to identify the main similarities (i.e. similar discussed topics) and differences between the countries.

Table 4.5 Main issues discussed by the FG participants in each of the seven countries

Germany (4 classes)	France (4 classes)	Greece (5 classes)	Spain (5 classes)	Poland (3 classes)	Romania (4 classes)	Estonia (6 classes)
Data Disclosure (20%)	Data Disclosure, Use & Re-gulation (16%)	Data disclosure & Use (21%)	Trust & Control (12%)	Control & Regulation (70%)	Privacy & relationships (15%)	Data Use & Risks (11%)
Protection & Respon-sibility (13%)	Respon-sibility (15%)	Data Use, Protection & Redress (24%)	Mandatory Disclosure & Regulation (11%)	Data Disclosure (22%)	Data Access & Consent (10%)	Anonymity (14%)
Conditions of Data Disclosure (23%)	Risks & Dangers (26%)	Data use & Consent (11%)	Social networking (9%)	Authenti-cation & Security (8%)	Monitoring & Regulation (22%)	Protection strategies (18%)
Security & Monitoring (44%)	Privacy invasion (43%)	Trust & Control (12%)	Virtuality (33%)	-	Data Disclosure (50%)	Experience & Trust (24%)
-	-	Identifi-cation (32%)	Monitoring & Privacy invasion (35%)	-	-	Public Data (19%)
-	-	-	-	-	-	Passwords (14%)

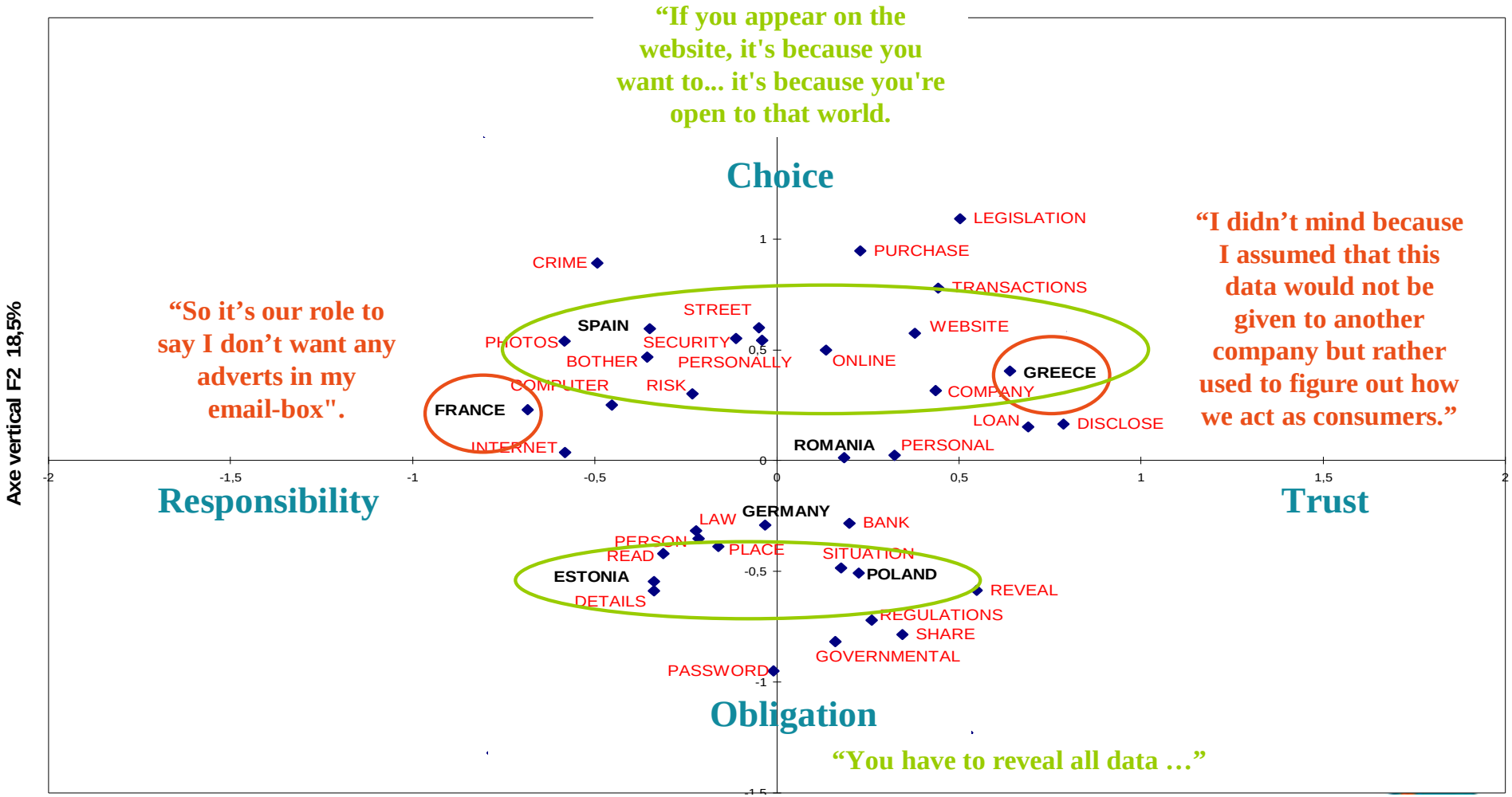
This comparative table highlights that:

- The issue of data disclosure has been discussed in all the countries; but in Estonia the discussion was more about public data (data that can be made public)
- The issues of control and protection have been discussed in all countries mainly to explain the importance of control and the absence of secure protection
- The issue of identification/authentication has mainly been discussed in Greece (unique identity document), in Poland (in association with security) and in Estonia
- The issue of risks is mainly present (as a topic) in France and Estonia
- The issue of monitoring (as a topic) has been mainly discussed in Spain and Romania
- The problem of data use has been largely debated in France, Romania and above all in Greece (where it appears in 3 topics)
- The issue of responsibility has mainly been discussed in Germany and France (i.e. the countries from the 'Old Europe' block)
- The issue of security has mainly been discussed in Germany (in relation to online payment) and in Poland (in relation to authentication)
- The issue of social networking has mainly been discussed in Spain and Romania

To reinforce the analysis of cultural differences, we now present the WordMapper© results. The similarities and differences between the seven countries appear on the graph below. It shows the most frequently employed words next to each country and notes the topics that mainly oppose the countries, topics illustrated by some participants' sentences just above.

To obtain this graph, the program produces, for each word and each variable value (i.e. each country) the absolute and relative contributions (to each axe) and the coordinates. The first two axes explain 20.27% and 18.50%, i.e. 38.77% of inertia. Romania and Germany are found in the middle of the graph because these countries are close to the average.

Figure 4.1 Factorial map per country: cultural similarities and differences in the 7 EU countries as regards to PIDM and privacy issues



The graph shows a clear opposition (on axe 1, so a main opposition) between France (Old Europe block, left part of the graph) and Greece (South block, right part). The results underline the importance of Responsibility in the case of France and the importance of Trust in the case of Greece. These results are coherent with the results obtained with Alceste© because responsibility already appears as an important topic for France (class number 2) with trust not appearing as a class and it is the opposite for Greece as responsibility doesn't appear as a class but trust appears as an topic (class number 4). There is thus an opposition between countries from the centre (France) and the South of Europe (Greece).

The factorial analysis per country also underlines that there is an opposition (on axe 2 so less important) between, on one side (top part of the graph) the two countries of the South block – Greece and Spain - and, on the other side, two countries from the East, one from the third block (Poland) and the other from the first block (Estonia). Axe 2 confirms the importance of trust for the two countries of the South Block (Greece and Spain) and the reluctance to give data to private entities for the two Eastern countries (Poland and Estonia). These results are coherent again with the results obtained with Alceste© which underline a proximity between Poland and Estonia on two topics (data disclosure and data use). So a clear difference appears between the Southern and the Eastern countries.

The results of Wordmapper© thus confirm the existence of a cultural divide in relation to PIDM. Although some similarities appear in people's perceptions about identity data disclosure and use, there is also some disparity between countries, especially as concerns:

- The perceived benefits in giving personal details and the trust one can have with the entities' use of personal data
- The control people have on their own data: is it possible to control future use of the data?
- The responsibility of data disclosure and (mis)use: who is responsible?
- The perceived efficiency of public regulation: does the existing data protection law really protect citizen's personal identity data against fraudulent use?
- The level of trust in public/private organisations
- The value and the acceptance of surveillance for security purpose

In addition to this cultural divide, we now study whether there is any generation divide.

4.4 Age analysis results: any generation divide?

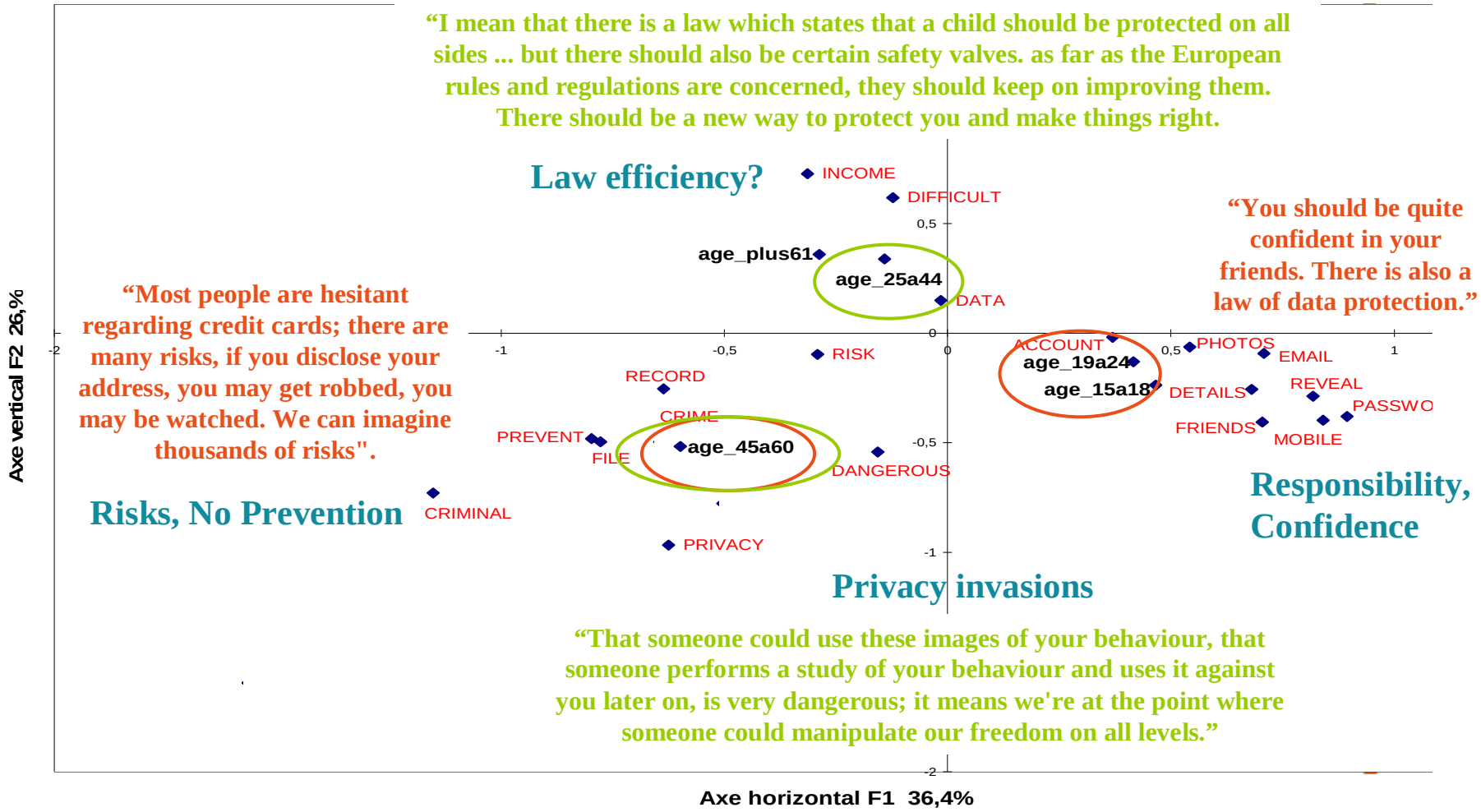
The similarities and differences between ages appear on the graph below. It shows the most frequently employed words next to each age category of the focus groups' participants and notes the topics that mainly oppose the young and adult generations, topics illustrated by some participants' sentences just above.

The plan of the first two axes explains 36.39% plus 26%, i.e. 62.39% inertia. There are four areas in the Factorial map but only three areas with the variable age. The first axe opposes the views of individuals aged 45 to 60 (left part of the graph) with those of young individuals who are less than 25 years old (right part). The second axe again opposes the views of individuals aged 45 to 60 (bottom part of the graph) with those of middle-aged (25 to 44 years old) individuals (top part).

Axe 1 opposes the views of older people (aged 45-60) who have relatively negative views with regards to PID disclosure and use. They think risks are numerous and they don't think that it is possible to prevent them. On the other side (right part of the graph), young people (aged 19-24) seem to be more positive, more responsible and to have some confidence in their ability to prevent possible data misuse. The same opposition appears again on axe 2 between, once more, older people (aged 45-60) who fear privacy invasions and mainly rely on public institutions and regulation to protect them and middle-aged people (aged 25-44) who have an intermediate position (not really confident but not untrusting neither) and who mainly discuss the efficiency of the data protection law.

We can therefore conclude that, in addition to the cultural divide, there is also clear generation divide with regards to PIDM and privacy issues, the young generation being more responsible and confident than adults.

Figure 4.2 Factorial map per age: generation similarities and differences in the 7 EU countries as regards to PIDM and privacy issues



5. Discussion and conclusion

With these results which have proven high convergence with both textual analysis methods used, a conclusion can now be made with implications, some limitations and perspectives of future work on this topic.

5.1 Implications

The research anticipates major scientific, methodological and policy implications.
From an academic point of view, the results first confirm most conclusions mentioned in the literature. Globally, the main topics discussed by the participants have already been studied by the main authors in the field, although not always together. Moreover, our results offer a profound description and understanding of all the issues of concern for the European citizens with regards to PIDM and privacy. They also permit the detailed study of the similarities and differences of people's attitudes and behaviours in relation to their age and country of origin. We not only confirm that these differences exist but also give details on which issues they are more particularly based upon and try to explain for what reasons.

However, several differences with previous studies have also been found, calling for a more profound analysis in future works. One important contribution to this work is the discussion about perceived responsibility in data handling. It appears that it is an important issue for people and that it could 1/ influence the choice of data protection strategies and 2/ explain the different paradoxes (e.g. the privacy paradox) with regards to consumers' behaviours. Another contribution concerns the perceived inefficiency of both data protection laws and other protection strategies. Our results show that many people feel powerless, helpless and want better protection; they value trust but do not know who they can really trust to protect their personal data. This calls for a more in-depth analysis in the future.

From a methodological point of view, this work is an application of qualitative research (taking the form of 14 focus groups) using triangulation as proof of the high reliability of the results. We used two different textual data analysis methods (i.e. content analysis and discriminate analysis) and two different software packages (i.e. Alceste© and Wordmapper©). The findings proved the complementarity of both methods and software and showed the high convergence of the results therefore confirming the importance of triangulation (e.g. Ammenwerth et al. 2003, Barnes & Vidgen 2006, Kaplan & Duchon 1988) to ensure the quality of qualitative data results.

Another exemplarity of this research lies in the elevated sample size and diversity, i.e. 7 EU countries, 139 participants from all gender, age, professional status and Internet level. This diversity offers the opportunity to test – and confirm - the existence of both cultural and generation divides.

In a policy point of view, the choice of discussion groups used in this qualitative research provided interesting information. This information is a tool with which to promote dialogue around various issues associated with personal identity data management and privacy concerns. By bringing together participants to discuss and challenge this information, such an approach is useful for understanding European public attitudes – at least those of the 7 countries surveyed - towards the management of personal identity data. With this study, it is possible to explore in depth, some key strands of debate in European and national policy. Recent reports and studies on the subject raise concerns about the privacy invasion of individuals. The results of this qualitative study allow the importance of these issues for European people to be determined and consequently consider whether a new regulatory framework is required.

5.2 Limitations

The most limiting factor of this study is linked with the - qualitative - method chosen which is mainly focused on motivations, risks and needs and don't really show the links between all of the interesting concepts. Additionally, the results are not really "generalisable" as the population interviewed remains relatively small in comparison with quantitative surveys.

5.3 Future work

Whereas this research offers a contribution to the knowledge of European citizens' disclosure practices, there is a need to understand in greater depth, people's perceptions and real behaviours as regards PIDM. This should help identifying ways to enhance their acceptation to disclose personal identity data and their adoption of new electronic identification and/or authentication means such as IPv6, electronic signature, biometrics... Further investigation is thus required to identify 1/ services that may improve secure disclosure at a minimum cost; 2/ solutions reducing the 'privacy paradox' and encouraging people to protect their identity data in more efficient ways and 3/ issues that need better regulatory protection.

As a first extension, a final questionnaire for a EU27 survey on this topic is being processed that has already been tested in 4 EU countries (Lusoli and Miltgen 2009) and should be conducted later in the year. This should help the European Commission direct the Information Society agenda in the years to come.

6. References

Ammenwerth, E., Iller, C. & Mansmann, U., 2003. Can evaluation studies benefit from triangulation? A case study. *International Journal of Medical Informatics*, 70(2-3), 237-248.

Aslanidou, S. & Menexes, G., 2008. Youth and the Internet: Uses and practices in the home. *Computers & Education*, 51(3), 1375-1391.

Barnes, S.J. & Vidgen, R.T., 2006. Data triangulation and web quality metrics: A case study in e-government. *Information & Management*, 43(6), 767-777.

Broos, A. & Roe, K., The digital divide in the playstation generation: Self-efficacy, locus of control and ICT adoption among adolescents. *Poetics*, 34(4-5), 306-317.

Campbell, A.J., 1997. Relationship marketing in consumer markets: A comparison of managerial and consumer attitudes about information privacy. *Journal of Direct Marketing*, 11(3), 44-57.

Changi Nam et al., 2006. Consumers' Privacy Concerns and Willingness to Provide Marketing-Related Personal Information Online. *Advances in Consumer Research*, 33(1), 212-217.

Dinev, T. & Hart, P., 2005. Internet Privacy Concerns and Social Awareness as Determinants of Intention to Transact. *International Journal of Electronic Commerce*, 10(2), 7-29.

Hoadley, C.M. et al., Privacy as information access and illusory control: The case of the Facebook News Feed privacy outcry. *Electronic Commerce Research and Applications*, 9(1), 50-60.

Hong Sheng, Fiona Fui-Hoon Nah & Keng Siau, 2008. An Experimental Study on Ubiquitous commerce Adoption: Impact of Personalization and Privacy Concerns. *Journal of the*

Association for Information Systems, 9(6), 344-376.

Kaplan, B. & Duchon, D., 1988. Combining qualitative and quantitative methods in information systems research: A case study. *MIS Quarterly: Management Information Systems*, 12(4), 571-586.

Lusoli, W. and Lancelot Miltgen, C.. (2009), Young people and emerging digital services: an exploratory surveys on motivations, perceptions and acceptance of risks, JRC Scientific and Technical Reports EUR 23765 EN, W. Lusoli, R. Compañó & I. Maghiros (Eds.), Sevilla: EC JRC IPTS

Lwin, M., Wirtz, J. & Williams, J.D., 2007. Consumer online privacy concerns and responses: a power–responsibility equilibrium perspective. *Journal of the Academy of Marketing Science*, 35(4), 572-585.

Lwin, M.O., Stanaland, A.J. & Miyazaki, A.D., 2008. Protecting children's privacy online: How parental mediation strategies affect website safeguard effectiveness. *Journal of Retailing*, 84(2), 205-217.

Milberg, S.J. et al., 1995. Values, personal information privacy, and regulatory approaches. *Commun. ACM*, 38(12), 65-74.

Montgomery, K., 2000. Youth and digital media: a policy research agenda. *Journal of Adolescent Health*, 27(2, Supplement 1), 61-68.

Papastergiou, M. & Solomonidou, C., 2005. Gender issues in Internet access and favourite Internet activities among Greek high school pupils inside and outside school. *Computers & Education*, 44(4), 377-393.

Phelps, J., Nowak, G. & Ferrell, E., 2000. Privacy Concerns and Consumer Willingness to Provide Personal Information. *Journal of Public Policy & Marketing*, 19(1), 27-41.

Tsai, M. & Tsai, C., 2010. Junior high school students' Internet usage and self-efficacy: A re-examination of the gender gap. *Computers & Education*, 54(4), 1182-1192.

Les autres documents de travail du GRANEM accessibles sur le site Web du laboratoire à l'adresse suivante : (http://ead.univ-angers.fr/~granem08/spip.php?rubrique36) :

Numéro	Titre	Auteur(s)	Discipline	Date
2008-01-001	The Cognitive consistency, the endowment effect and the preference reversal phenomenon	Serge Blondel, Louis Lévy-Garboua	Théorie du Risque	octobre 2008
2008-02-002	Volatility transmission and volatility impulse response functions in European electricity forward markets	Yannick Le Pen, Benoît Sévi	Econométrie Appliquée	octobre 2008
2008-03-003	Anomalies et paradoxes dans le cas des choix alimentaires : et si les carottes n'étaient pas oranges ?	Serge Blondel, Christophe Daniel, Mahsa Javaheri	Economie Expérimentale	octobre 2008
2008-04-004	The effects of spatial spillovers on the provision of urban environmental amenities	Johanna Choumert, Walid Oueslati, Julien Salanié	Economie du Paysage	octobre 2008
2008-05-005	Why do rational people vote in large elections with costs to vote?	Serge Blondel, Louis Lévy-Garboua	Théorie du Risque	novembre 2008
2008-06-006	Salaires, conditions et satisfaction au travail	Christophe Daniel	Economie du Travail	novembre 2008
2008-07-007	Construction communicationnelle du stock de connaissances de la compétence collective – Contribution à partir d'une conversation.	Nicolas Arnaud	Gestion des Ressources Humaines	décembre 2008
2008-08-008	On the non-convergence of energy intensities: evidence from a pair-wise econometric approach	Yannick Le Pen, Benoît Sévi	Econométrie Appliquée	décembre 2008
2008-09-009	Production of Business Ethics	Guido Hülsmann	Economie Politique	décembre 2008
2008-10-010	Time preference and investment expenditure	Guido Hülsmann	Economie Politique	décembre 2008
2008-11-011	Le marché de la photographie contemporaine est-il soluble dans celui de l'art contemporain ?	Dominique Sagot-Duvauroux	Economie de la Culture	décembre 2008
2008-12-012	The newsvendor problem under multiplicative background risk	Benoît Sévi	Microéconomie de l'Incertain	décembre 2008
2009-01-013	Complémentarité de la collaboration électronique et de l'investissement relationnel : étude de cas exploratoire d'un SIIO dans le secteur du meuble	Redouane Elamrani, Nicolas Arnaud	Organisation	avril 2009
2009-02-014	On the realized volatility of the ECX CO_2 emissions 2008 futures contract: distribution, dynamics and forecasting	Julien Chevallier, Benoît Sévi	Finance	mai 2009
2009-03-015	The communicational making of a relation-specific skill: contributions based on the analysis of a conversation to strategy-as-practice and resource-based view perspectives	Nicolas Arnaud	Stratégie	juin 2009
2009-04-016	Le droit d'auteur, incitation à la création ou frein à la diffusion ? Une analyse empirique du cas de la création télévisuelle	Françoise Benhamou, Stéphanie Peltier	Economie de la Culture	septembre 2009
2009-05-017	Diversity analysis in cultural economics: theoretical and empirical considerations	Françoise Benhamou, Renato G. Flôres Jr., Stéphanie Peltier	Economie de la Culture	septembre 2009
2009-06-18	L'épargne retraite en entreprise : un état des lieux au regard de l'expérience américaine	Fabrice Pansard, Bruno Séjourné	Finance	septembre 2009
2009-07-19	Options introduction and volatility in the EU ETS	Julien Chevallier, Yannick Le Pen, Benoît Sévi	Econométrie Appliquée	septembre 2009
2009-08-20	Modeling strategic interactions between firms and local authorities – The case of a biotechnology cluster	Alain Berro, Isabelle Leroux	Economie des réseaux	septembre 2009
2009-09-21	The strategy adopted by non-profit care services organizations in dealing with the new French regulatory system: strategic coalitions and reterritorialisation of activities	Isabelle Leroux, Laurent Pujol, Eric Rigamonti	Economie Sociale	novembre 2009
2009-10-22	Une nouvelle lecture du territoire par la limite	Jean-Claude Taddei	Territoire	novembre 2009
2010-01-23	Adoption of new identity-based services: Proposition of a conceptual model based on TAM, DOI and perceived risks	Caroline Lancelot Miltgen	e-marketing	juillet 2010
2010-02-24	Young Europeans' motivations, perceived risks and requirements regarding electronic identification : Some comparative results from focus groups in four EU27 countries	Caroline Lancelot Miltgen	e-marketing	décembre 2010
2010-03-25	Analyse du risque de non-exécution des ordres à la bourse de Paris	Angélique Aubier Piron	Finance	décembre 2010
2011-01-26	Who cares? Europeans' attitudes towards the disclosure of personal identity data	Caroline Lancelot Miltgen, Margherita Bacigalupo, Wainer Lusoli	Systèmes d'information et e-marketing	janvier 2011